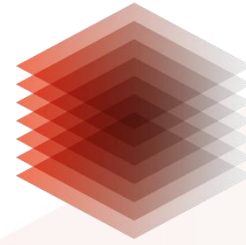
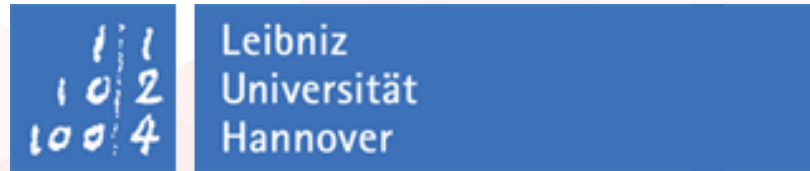


LEIBNIZ INFORMATION CENTRE
FOR SCIENCE AND TECHNOLOGY
UNIVERSITY LIBRARY



TIB



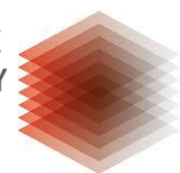
Von Open Access zu Open Knowledge - wie wir Informationsflüsse der Wissenschaft in der Digitalen Welt organisieren können

Prof. Dr. Sören Auer
Leibniz University of Hannover
TIB Technische Informationsbibliothek

11
102
1004

Leibniz
Universität
Hannover

LEIBNIZ INFORMATION CENTRE
FOR SCIENCE AND TECHNOLOGY
UNIVERSITY LIBRARY



TIB



Gottfried Wilhelm Leibniz

* 21. Juni/ 1. Juli 1646 in Leipzig

† 14. November 1716 in Hannover



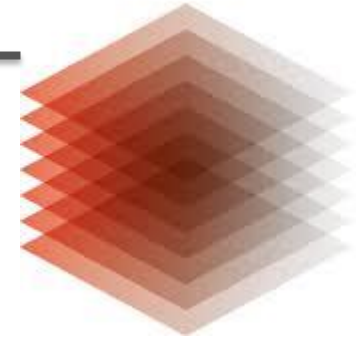
Namesake

Namesake



Member of

Library of



TIB



How did information flows change in the digital era?



Computer



Commodore C64

Abb. 1-6

GLEICH ZUM MITNEHMEN!

In Ihrem OTTO wohnen Einrichtungshaus. Mehr darüber Seite 1026.

1 Farbmonitor 598.-

2 Grün-Monitor 198.-

3 Datasette 99.-

4 »C 64 C« 448.-

5 »Maus« 129.-

6 Joystick 19.90

7 Floppy »VC 1541« ab 448.-

8 5 1/4-Disketten

9 Diskettenkasten 29.90

10 39.90

11 ab 39.95

12 59.90

13 49.98

14 ab 39.95

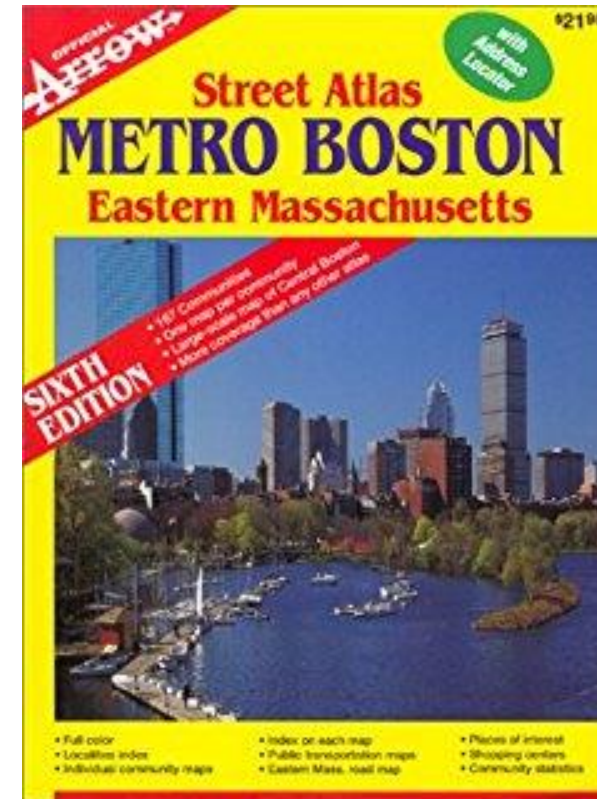
15-17 49.-

GEOS

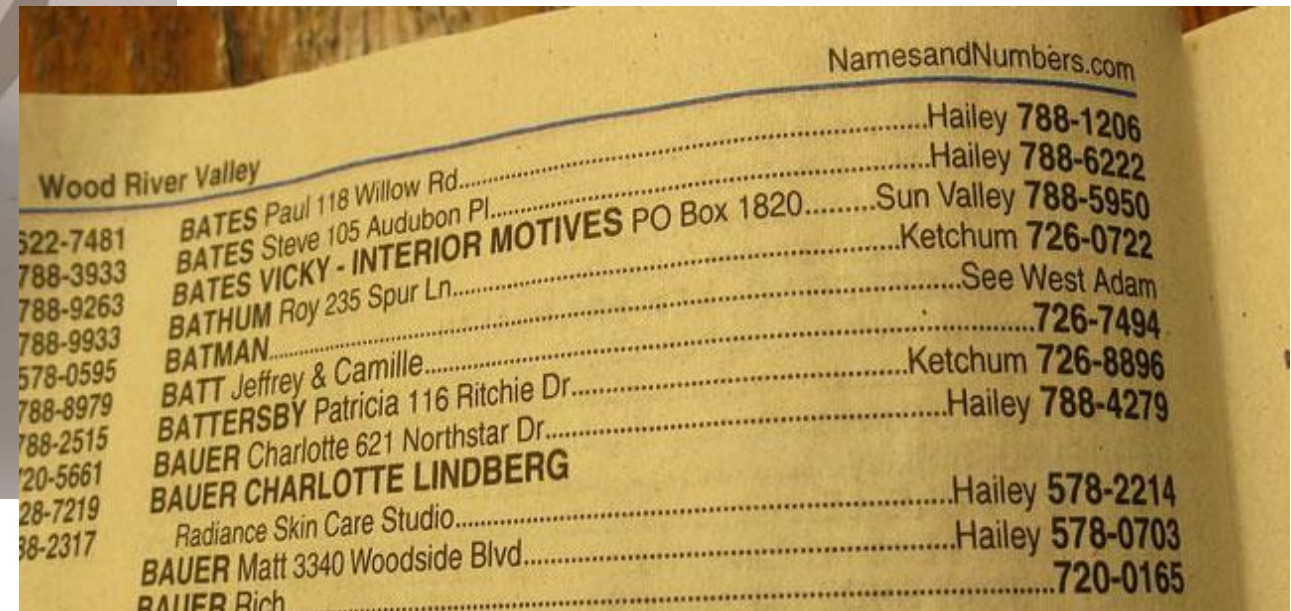
BASF NewDisk

SANYO

Road Maps



Phone Books



**How does it
work today?**

Smartphones (4.691)*

Beliebteste Produkttypen

LTE Smartphone
3.625 ProduktePhablet
2.023 ProdukteAndroid Smartphone
4.145 Produkte

Preis

<input type="checkbox"/> bis 95 €	900
<input type="checkbox"/> 95 € bis 150 €	1.185
<input type="checkbox"/> 150 € bis 200 €	857
<input type="checkbox"/> 200 € bis 300 €	840
<input type="checkbox"/> ab 300 €	909

Serie

<input type="checkbox"/> Apple iPhone	223
<input type="checkbox"/> Samsung Galaxy S	216
<input type="checkbox"/> Samsung Galaxy A	95
<input type="checkbox"/> Huawei P	165
<input type="checkbox"/> Huawei Mate	58

[+ mehr](#)Ergebnisse in **Smartphones** werden angezeigt. Ergebnisse in allen Kategorien anzeigen.

Samsung Galaxy S10

LTE Smartphone, 6,1 Zoll, Quad HD,
8 GB RAM, iOS 9.0 Pie, 16 Mega-
pixel, 3.400 mAh, Speicher ...mehr

Note Ø 1,5 ★★★★★ 8

388 Angebote

664,99 – 1.341,40 €



Apple iPhone Xr

LTE Smartphone, 6,1 Zoll, HD,
3 GB RAM, iOS 12, 12 Megapixel,
2.942 mAh, Gewicht 194 g

Note Ø 1,6 ★★★★★ 111

923 Angebote

695,00 – 1.284,01 €



Huawei P30 Pro

LTE Smartphone, 6,47 Zoll, Full HD,
Android 9.0 Pie, 40 Megapixel,
4.200 mAh, Speicher erwe ...mehr

Note Ø 1,5 ★★★★★ 2

196 Angebote

739,00 – 1.391,11 €



Xiaomi Mi 9

LTE Smartphone, 6,39 Zoll, Full HD,
6 GB RAM, Android 9.0 Pie, 48 Mega-
pixel, 3.300 mAh, Gewicht 173 g

Note Ø 2,0 ★★★★★ 4

52 Angebote

445,00 – 580,00 €



Samsung Galaxy S10e

LTE Smartphone, 5,8 Zoll, Full HD,
Android 9.0 Pie, 16 Megapixel,
3.100 mAh, Speicher erwe ...mehr

Note Ø 1,8 ★★★★★ 2

294 Angebote

549,49 – 857,79 €



Apple iPhone Xs

LTE Smartphone, 5,8 Zoll, Full HD,
4 GB RAM, iOS 12, 12 Megapixel,
2.658 mAh, Gewicht 177 g

Note Ø 1,6 ★★★★★ 235

482 Angebote

895,00 – 1.902,81 €



Huawei P20

LTE Smartphone, 5,8 Zoll, Full HD,
4 GB RAM, Android 9.0 Pie, 20 Mega-
pixel, 3.400 mAh, Gewicht 165 g

Note Ø 1,8 ★★★★★ 222

194 Angebote

340,00 – 649,00 €



Samsung Galaxy S10 Plus

LTE Smartphone, 6,4 Zoll, Quad HD,
Android 9.0 Pie, 16 Megapixel,
4.100 mAh, Speicher erweiterbar

Note Ø 1,6 ★★★★★ 3

247 Angebote

859,99 – 1.940,34 €

Anzeige

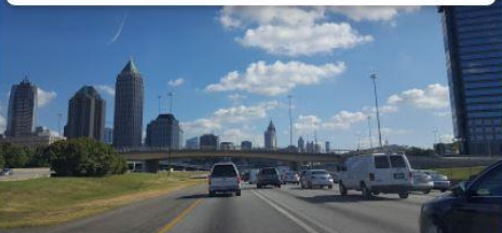


SIEMENS

TESTSIEGER
Stiftung
Warentest
SEHR GUT
(1,2)

Der Testsieger
für optimale
Frische.

Jetzt entdecken! >



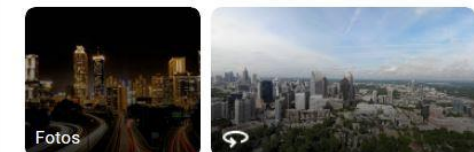
Atlanta

Georgia
USA

Klar und vereinzelt Wolken · 17 °C
06:02

Routenplaner Speichern In der Nähe An mein Smartphone senden Teilen

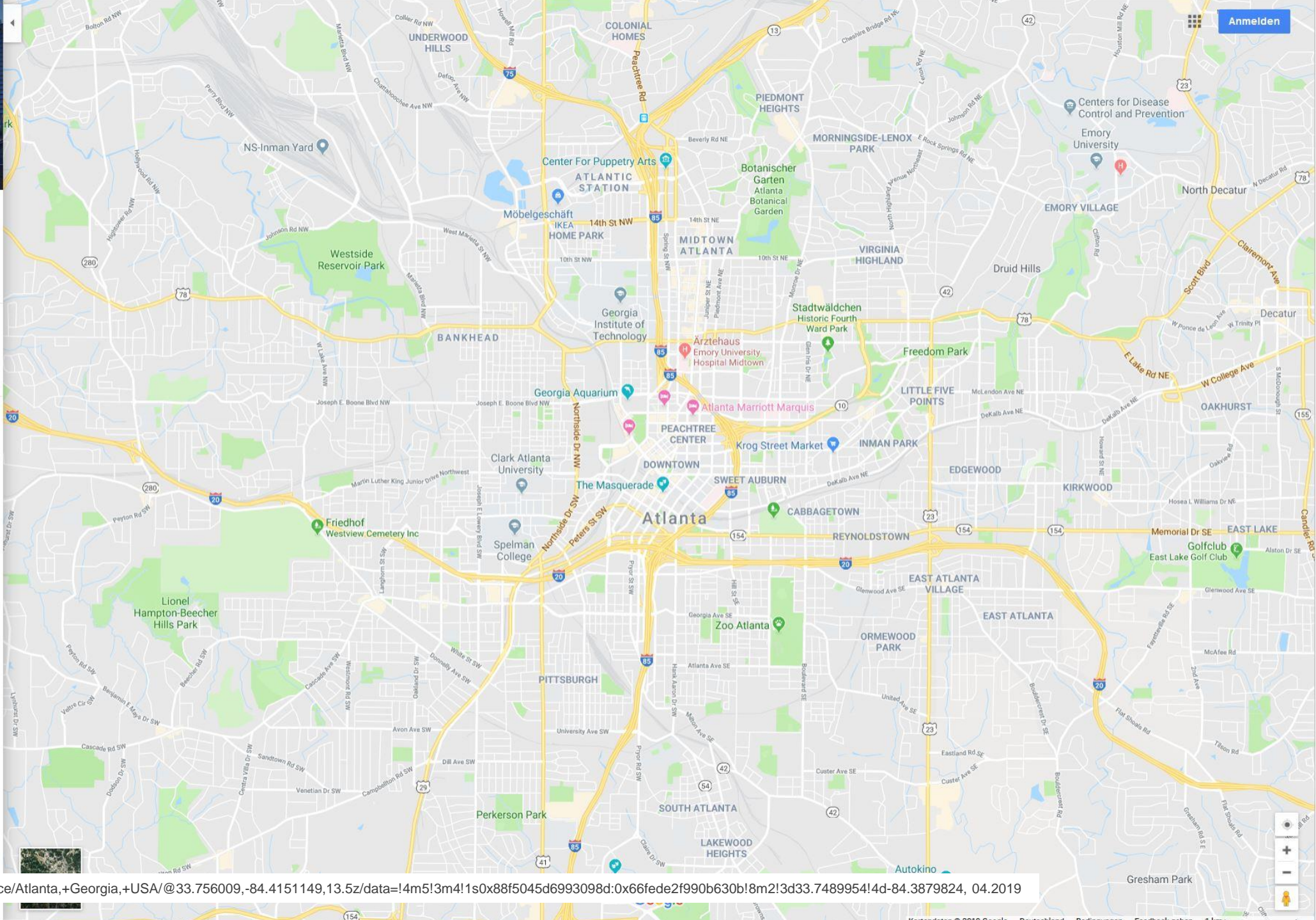
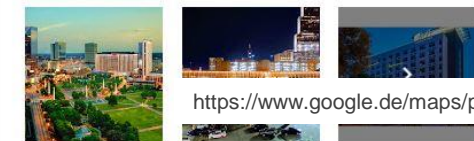
Fotos



Kurzinfo

Atlanta ist die Hauptstadt des US-Bundesstaates Georgia. Die Stadt spielte sowohl während des Amerikanischen Bürgerkriegs als auch während der Bürgerrechtsbewegung in den 1960er-Jahren eine wichtige Rolle. Das Atlanta History Center erzählt von der Geschichte der Stadt, während die Martin Luther King Jr. National Historic Site dem Leben und Wirken des afroamerikanischen Bürgerrechtlers gewidmet ist. Im Centennial Olympic Park, der für die Olympischen Spiele 1996 in Downtown gebaut wurde, befindet sich das eindrucksvolle Georgia Aquarium.

Hotels



The World of Publishing & Communication has profoundly changed

- **New means adapted to the new possibilities** were developed, e.g. „zooming“, dynamics
- **Business models** changed completely
- More focus on data, interlinking of **data / services and search** in the data
- Integration, **crowdsourcing** play an important role

What about Scholarly Communication?

Scientific publishing in the 17th century

One of the earliest research journals:
Philosophical Transactions of the Royal Society

13

PHILOSOPHICAL
TRANSACTIONS:
GIVING SOME
ACCOMPT
OF THE PRESENT
Undertakings, Studies, and Labours
OF THE
INGENIOUS
IN MANY
CONSIDERABLE PARTS
OF THE
WORLD.

Vol I.

For *Anno* 1665, and 1666.

In the SAVOY,
Printed by T. N. for John Martyn at the Bell, a little with-
out Temple-Bar, and James Allestry in Duck-Lane,
Printers to the Royal Society.

Presented by the Author May 30th 1667.

Scholarly communication in 1865

THE INTELLECTUAL OBSERVER.

JANUARY, 1865.

CELESTIAL CHEMISTRY, AND THE PHYSICAL CONSTITUTION OF THE STARS AND NEBULÆ.

BY THOMAS W. BURE, F.R.A.S., F.C.S.

(With a Coloured Plate.)

FEW things are more remarkable in the present aspect of science than the manner in which its various departments come into contact one with another, thus aiding the student in a way quite unlooked for, and throwing light upon the subject of research from a quarter whence it was least expected. As when stones are thrown into water, so the circle of each science at first seems to be totally distinct from all the others, but gradually these separate circles enlarge and widen, until they intersect and produce larger circles and wider generalizations in the increasing domain of human knowledge. Thus, chemistry was, in the time of Davy, furnished with a new and powerful analytical agent in the shape of voltaic electricity, and the same agency, which is itself evoked by chemical action, has given us the long series of discoveries in electro magnetism, culminating in the splendid practical application of the electric telegraph. So, too, photography, which is essentially chemical in its nature, has been of the greatest service to the physicist in furnishing him with a constant and unerring record of the indications of his barometer, thermometer, and magnetic instru-

Publishing in 1970s

A Relational Model of Data for Large Shared Data Banks

E. F. CODD

IBM Research Laboratory, San Jose, California

Future users of large data banks must be protected from having to know how the data is organized in the machine (the internal representation). A prompting service which supplies such information is not a satisfactory solution. Activities of users at terminals and most application programs should remain unaffected when the internal representation of data is changed and even when some aspects of the external representation are changed. Changes in data representation will often be needed as a result of changes in query, update, and report traffic and natural growth in the types of stored information.

Existing noninferential, formatted data systems provide users with tree-structured files or slightly more general network models of the data. In Section 1, inadequacies of these models are discussed. A model based on n -ary relations, a normal form for data base relations, and the concept of a universal data sublanguage are introduced. In Section 2, certain operations on relations (other than logical inference) are discussed and applied to the problems of redundancy and consistency in the user's model.

KEY WORDS AND PHRASES: data bank, data base, data structure, data organization, hierarchies of data, networks of data, relations, derivability, redundancy, consistency, composition, join, retrieval language, predicate calculus, security, data integrity

CR CATEGORIES: 3.70, 3.73, 3.75, 4.20, 4.22, 4.29

The relational view (or model) of data described in Section 1 appears to be superior in several respects to the graph or network model [3, 4] presently in vogue for non-inferential systems. It provides a means of describing data with its natural structure only—that is, without superimposing any additional structure for machine representation purposes. Accordingly, it provides a basis for a high level data language which will yield maximal independence between programs on the one hand and machine representation and organization of data on the other.

A further advantage of the relational view is that it forms a sound basis for treating derivability, redundancy, and consistency of relations—these are discussed in Section 2. The network model, on the other hand, has spawned a number of confusions, not the least of which is mistaking the derivation of connections for the derivation of relations (see remarks in Section 2 on the “connection trap”).

Finally, the relational view permits a clearer evaluation of the scope and logical limitations of present formatted data systems, and also the relative merits (from a logical standpoint) of competing representations of data within a single system. Examples of this clearer perspective are cited in various parts of this paper. Implementations of systems to support the relational model are not discussed.

1.2. DATA DEPENDENCIES IN PRESENT SYSTEMS

The provision of data description tables in recently developed information systems represents a major advance toward the goal of data independence [5, 6, 7]. Such tables facilitate changing certain characteristics of the data representation stored in a data bank. However, the variety of data representation characteristics which can be changed *without logically impairing some application programs* is still quite limited. Further, the model of data with which users interact is still cluttered with representational prop-

Scientific publishing today

WE HAVE



BUT

- Mainly based on PDF
- Is only partially machine-readable
- Does not preserve structure
- Does not allow embedding of semantics
- Does not facilitate interactivity / dynamicity / repurposing
- ...

AGDISTIS - Graph-Based Disambiguation of Named Entities using Linked Data

Ricardo Usbeck^{1,2}, Axel-Cyrille Ngonga Ngomo¹, Michael Röder^{1,2}, Daniel Gerber¹, Sandro Athaide Coelho³, Sören Auer⁴, and Andreas Both²

¹ University of Leipzig, Germany, ² R & D, Unister GmbH, Germany, ³ Federal University of Juiz de Fora, Brazil, ⁴ University of Bonn & Fraunhofer IAIS, Germany
email: {usbeck|ngonga}@informatik.uni-leipzig.de

Abstract. Over the last decades, several billion Web pages have been made available on the Web. The ongoing transition from the current Web of unstructured data to the Web of Data yet requires scalable and accurate approaches for the extraction of structured data in RDF (Resource Description Framework) from these websites. One of the key steps towards extracting RDF from text is the disambiguation of named entities. While several approaches aim to tackle this problem, they still achieve poor accuracy. We address this drawback by presenting AGDISTIS, a novel knowledge-base-agnostic approach for named entity disambiguation. Our approach combines the Hypertext-Induced Topic Search (HITS) algorithm with label expansion strategies and string similarity measures. Based on this combination, AGDISTIS can efficiently detect the correct URIs for a given set of named entities within an input text. We evaluate our approach on eight different datasets against state-of-the-art named entity disambiguation frameworks. Our results indicate that we outperform the state-of-the-art approach by up to 29% F-measure.

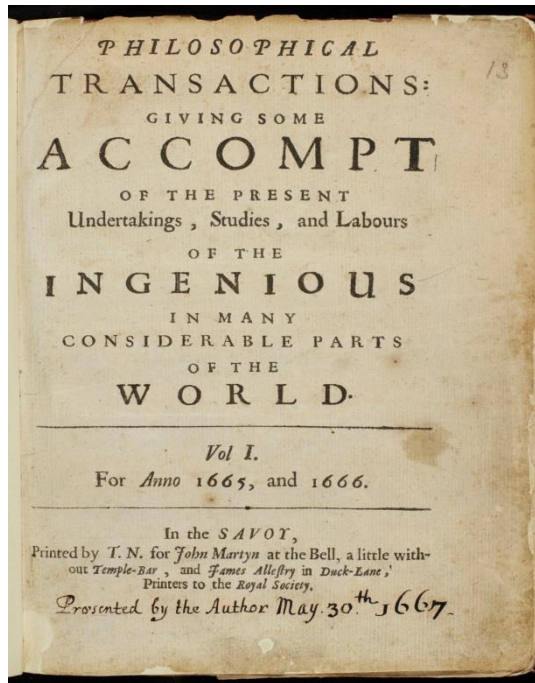
1 Introduction

The vision behind the Web of Data is to provide a new machine-readable layer to the Web where the content of Web pages is annotated with structured data (e.g., RDFa [1]). However, the Web in its current form is made up of at least 15 billion Web pages.¹ Most of these websites are unstructured in nature. Realizing the vision of a usable and up-to-date Web of Data thus requires scalable and accurate natural-language-processing approaches that allow extracting RDF from such unstructured data. Three tasks play a central role when extracting RDF from unstructured data: named entity recognition (NER), named entity disambiguation (NED), also known as entity linking [16], and relation extraction (RE). For the first sentence of Example 1, an accurate named entity recognition approach would return the strings *Barack Obama* and *Washington, D.C.*. A high-quality DBpedia-based named entity disambiguation (NED) approach would use these already recognized named entities and map the strings

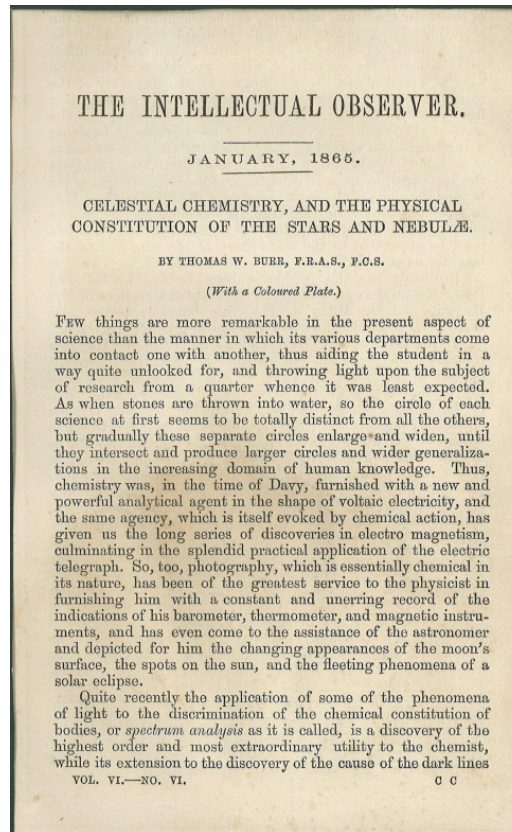
¹ Data gathered from <http://www.worldwidewebsize.com/> on January 4th, 2014.

Scholarly Communication has not changed (much)

17th century



19th century



20th century



21st century



Meanwhile other information intense domains were completely disrupted: mail order catalogs, street maps, phone books, ...

We need to rethink the way how research is represented and communicated

Challenges we are facing:

Digitalisation of Science

- Data integration and analysis
- Digital collaboration

Monopolisation by commercial actors

- Publisher look-in effects
- Maximization of profits ^[1]

Reproducibility Crisis

- Majority of experiments are hard or not reproducible ^[2]

Proliferation of publications

- Publication output doubled within a decade
- continues to rise ^[3]

Deficiency of Peer Review

- Deteriorating quality ^[4]
- Predatory publishing

[1] <http://thecostofknowledge.com>, <https://www.projekt-deal.de>

[2] M. Baker: *1,500 scientists lift the lid on reproducibility*, *Nature*, 2016.

[3] *Science and Engineering Publication Output Trends*, National Science Foundation, 2018.

[4] J. Couzin-Frankel: *Secretive and Subjective, Peer Review Proves Resistant to Study*, *Science*, 2013.

Proliferation of scientific literature

Science and engineering articles by region, country: 2004 and 2014

Rank	Region, country, or economy	2004	2014	Average annual growth rate (%)	2014 world total (%)	2014 cumulative world total (%)
na	World	1,272,362	2,290,294	6.1	100.0	na
1	United States	336,194	431,623	2.5	18.8	18.8
2	China	110,388	395,588	13.6	17.3	36.1
3	Germany	72,177	107,747	4.1	4.7	40.8
4	India	28,752	106,574	14.0	4.7	45.5
5	Japan	95,999	103,793	0.8	4.5	50.0
6	United Kingdom	75,119	101,536	3.1	4.4	54.4
7	France	53,375	74,269	3.4	3.2	57.7
8	Italy	42,647	70,453	5.1	3.1	60.8
9	South Korea	27,029	63,748	9.0	2.8	63.5
10	Canada	40,624	60,916	4.1	2.7	66.2
11	Spain	30,977	56,604	6.2	2.5	68.7
12	Brazil	18,814	53,152	10.9	2.3	71.0
13	Australia	26,277	52,269	7.1	2.3	73.3
14	Russia	26,869	43,487	4.9	1.9	75.2
15	Iran	4,952	36,539	22.1	1.6	76.8

Reproducibility Crisis

1,500 scientists lift the lid on reproducibility

Monya Baker in *Nature*, 2016. **533** (7604): 452–454.
[doi:10.1038/533452a](https://doi.org/10.1038/533452a):

- 70% failed to reproduce at least one other scientist's experiment
- 50% failed to reproduce one of their own experiments

Failure to reproduce results among disciplines (in brackets own results)

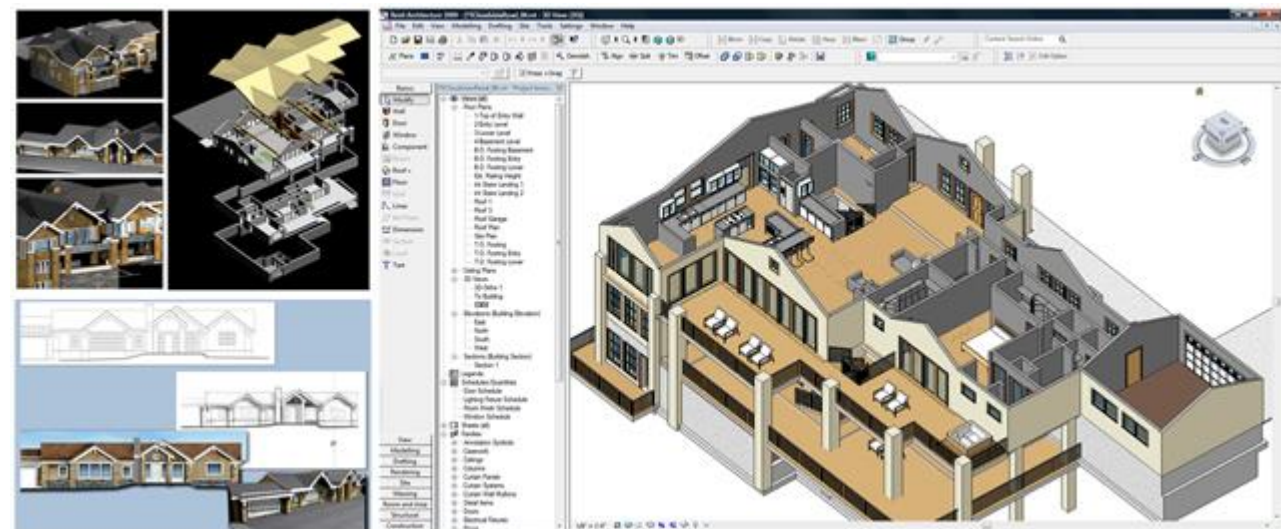
chemistry	87%	(64%)
biology	77%	(60%)
physics and engineering	69%	(51%)
Earth sciences	64%	(41%)



Duplication and Inefficiency

How can we avoid duplication if the terminology, research problems, approaches, methods, characteristics, evaluations, ... are not properly defined and identified?

How would you build an engine / building without properly defining their parts, relationships, materials, characteristics ...?



Root Cause – Deficiency of Scholarly Communication?

Lack of...

Transparency

information is hidden
in text

Integratability

fitting different
research results
together

Machine assistance

unstructured content
is hard to process

Identifiability

of concepts beyond
metadata

Collaboration

one brain barrier

Overview

Schientists look for
the needle in the
haystack

CRISPR

☐ Nur im Bibliothekskatalog der TIB suchen

**Search for CRISPR:
> 9.000 Results**

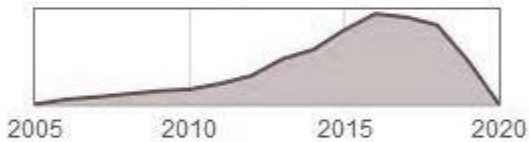
[Hier geht es zum klassischen Katalog ↗](#)

Seite 1 von 9.881 Ergebnissen

Sortieren nach: Relevanz | [Aktualität](#) | [Titel](#)

Treffer filtern

Erscheinungsjahr

 -

Medientyp

- ☐ Aufsatz (Zeitschrift) (6.870)
- ☐ Patent (2.619)
- ☐ Hochschulschrift (127)
- ☐ Sonstige (82)
- ☐ Aufsatz (Konferenz) (58)

[+ Mehr anzeigen](#)Für folgende Begriffe wurden Synonyme verwendet: [CRISPR](#)▼Suche ohne Synonyme: [CRISPR](#)

« < 1 2 3 4 5 6 7 > »



CRISPR human trial.(clustered regularly interspaced short palindromic repeats)

Online Contents | 2016

cited: 0 | score: 310.36673



Advances in clustered regularly interspaced short palindromic repeats - A review

Wang, L. / He, J. / Wang, J. | British Library Online Contents | 2011

cited: 0 | score: 297.1446



Google Scholar

CRISPR

Artikel

Ungefähr 238.000 Ergebnisse (0,04 Sek.)

Beliebige Zeit
Seit 2019
Seit 2018
Seit 2015
Zeitraum wählen

Nach Relevanz sortieren
Nach Datum sortieren

Beliebige Sprache
Seiten auf Deutsch

☒ Patente einschließen
☒ Zitate einschließen

☒ Alert erstellen

Multiplex genome engineering using **CRISPR**/Cas systems
Cong, FA Ran, D Cox, S Lin, R Barretto... - ..., 2013 - science.sciencemag.org
Functional elucidation of causal genetic variants and elements requires precise genome editing using CRISPR-Cas systems

**Search for CRISPR:
> 238.000 Results**

**How good is CRISPR
(wrt. precision, safety, cost)?**

**What specifics has genome
editing with insects?**

**Who has applied it to
butterflies?**

CRISPR provides acquired immunity to bacteria and archaea by using clustered regularly interspaced short palindromic repeats (CRISPR) adaptive immune system can be used to facilitate efficient genome editing

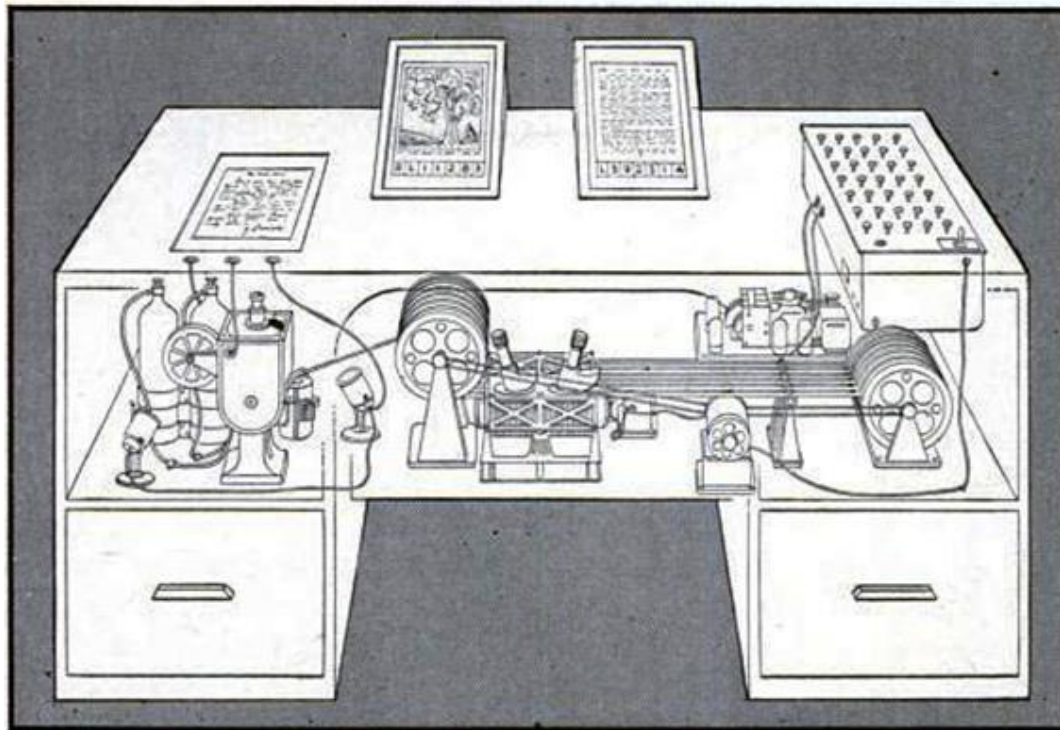
Genome engineering using CRISPR/Cas systems
FA Ran, PD Hsu, J Wright, V An, S Lin, D Cox, S Lin, R Barretto... - ..., 2013 - nature.com
Targeted nucleases are powerful tools for mediating genome alteration with high precision. The RNA-guided Cas9 nuclease from the microbial clustered regularly interspaced short palindromic repeats (CRISPR) adaptive immune system can be used to facilitate efficient genome editing

☆ ⓘ Zitiert von: 3418 Ähnliche Artikel Alle 19 Versionen Web of Science: 2423

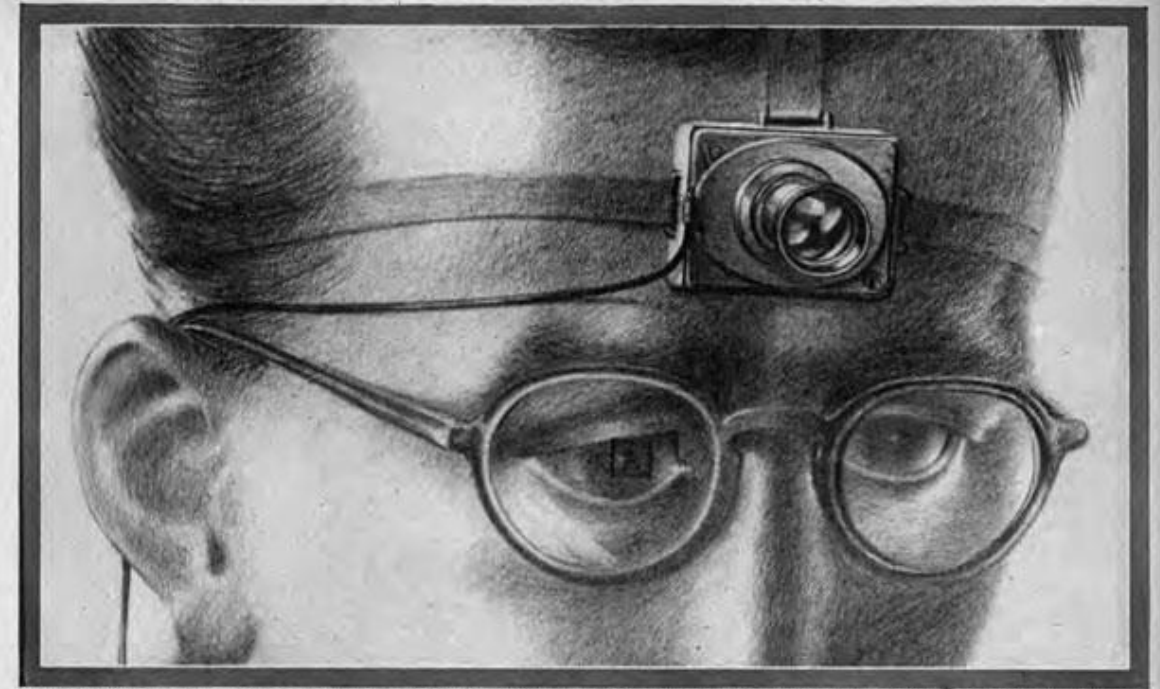
☆ ⓘ Zitiert von: 3650 Ähnliche Artikel Alle 19 Versionen Web of Science: 2423

**How can
we fix it?**

Realizing Vannevar Bush's vision of Memex



Source: <http://photos1.blogger.com/blogger/5874/1071/1600/Memex.jpg>



A SCIENTIST OF THE FUTURE RECORDS EXPERIMENTS WITH A TINY CAMERA FITTED WITH UNIVERSAL-FOCUS LENS. THE SMALL SQUARE IN THE EYEGLASS AT THE LEFT SIGHTS THE OBJECT

AS WE MAY THINK

A TOP U. S. SCIENTIST FORESEES A POSSIBLE FUTURE WORLD
IN WHICH MAN-MADE MACHINES WILL START TO THINK

by VANNEVAR BUSH

DIRECTOR OF THE OFFICE OF SCIENTIFIC RESEARCH AND DEVELOPMENT
Condensed from the *Affonic Monthly*, July 1945

This has not been a scientists' war; it has been a war in which all have had a part. The scientists, burying their old professional competition in the demand of a common cause, have shared greatly and learned much. It has been exhilarating to work in effective partnership. What are the scientists to do next?

For the biologists, and particularly for the medical scientists, there can be little indecision, for their war work has hardly required them to leave the old paths. Many indeed have been able to carry on their war research in their familiar peacetime laboratories. Their objectives remain much the same.

It is the physicists who have been thrown most violently off stride, who have left academic pursuits for the making of strange destructive gadgets, who have had to devise new methods for their unanticipated assignments. These have done their best on the deficit that made it possible to win the

war, and the effort to bridge between disciplines is correspondingly superficial.

Professionally our methods of transmitting and reviewing the results of research are generations old and by now are totally inadequate for their purpose. If the aggregate time spent in writing scholarly works and in reading them could be evaluated, the ratio between these amounts of time might well be startling. Those who conscientiously attempt to keep abreast of current thought, even in restricted fields, by close and continuous reading might well shy away from an examination calculated to show how much of the previous month's efforts could be produced on call.

Mendel's concept of the laws of genetics was lost to the world for a generation because his publication did not reach the few who were capable of understanding it. The sort of catastrophe is undoubtedly being repeated as significant attainments become lost in the mass

been part of a great team. Now one asks where they will find objectives

Publication has been extended far beyond our present ability to make real

Source: <http://tntindex.blogspot.com/2014/10/tabletalk-vannevar-bushs-memex.html>

Concepts

Overarching Concepts

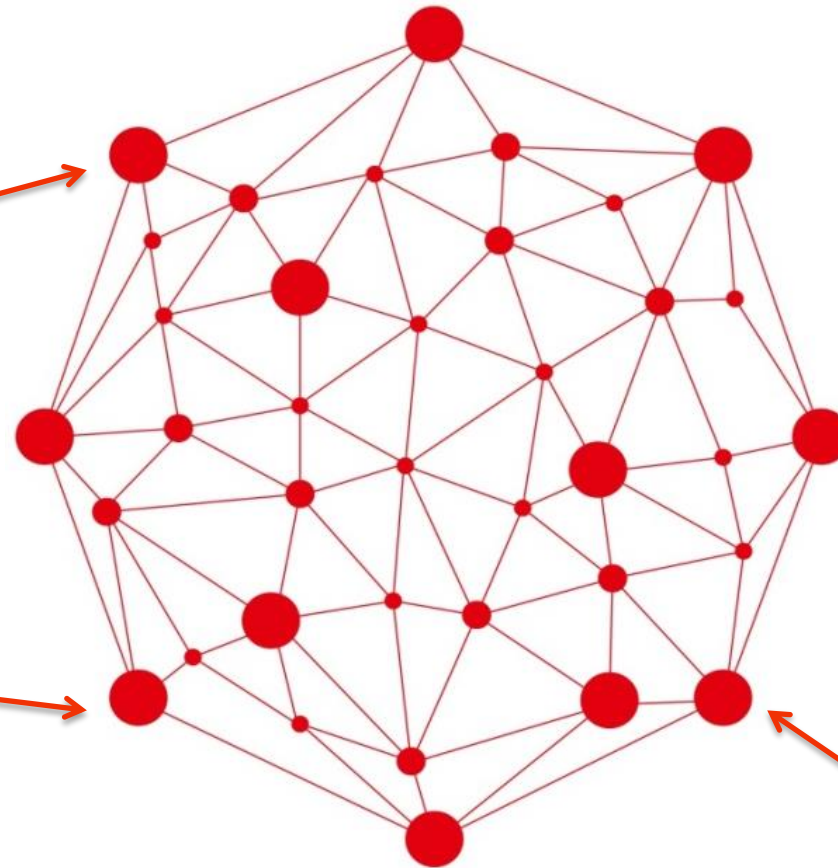
- Research problems
- Definitions
- Research approaches
- Methods

Artefacts

- Publications
- Data
- Software
- Image/Audio/Video
- Knowledge Graphs / Ontologies

Domain specific Concepts

Mathematics	Physics	Chemistry	Computer Science	Technology	Architecture
<ul style="list-style-type: none"> • Definitions • Theorems • Proofs • Methods • ... 	<ul style="list-style-type: none"> • Experiments • Data • Models • ... 	<ul style="list-style-type: none"> • Substances • Structures • Reactions • ... 	<ul style="list-style-type: none"> • Concepts • Implementations • Evaluations • ... 	<ul style="list-style-type: none"> • Standards • Processes • Elements • Units, Sensor data 	<ul style="list-style-type: none"> • Regulations • Elements • Models • ...



Linked Data Principles



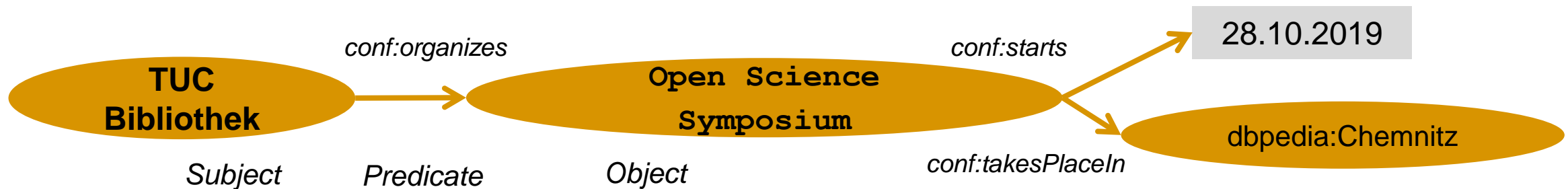
Addressing the neglected third V (Variety)

1. Use **URIs** to **identify** the “things” in your data
2. Use **http:// URIs** so people (and machines) can **look** them **up** on the web
3. When a URI is looked up, **return** a **description** of the thing **in** the W3C **Resource Description Format (RDF)**
4. Include **links to related things**

<http://www.w3.org/DesignIssues/LinkedData.html>

RDF & Linked Data in a Nutshell

1. Graph based RDF data model consisting of S-P-O statements (facts)

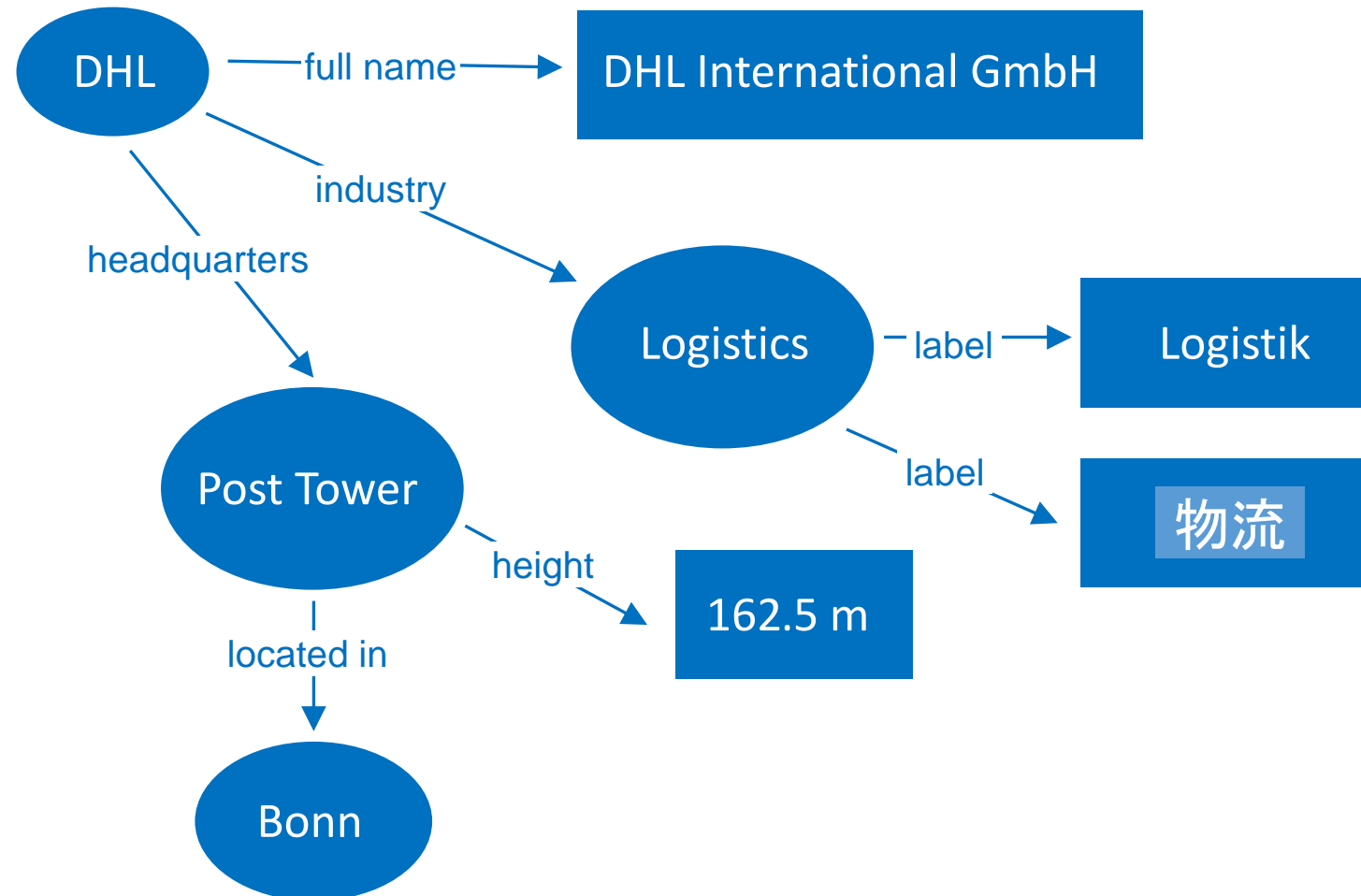


2. Serialised as RDF Triples:

TUC_Bibliothek	conf:organizes	OpenScienceSymposium .
OpenScienceSymposium	conf:starts	"2019-10-28"^^xsd:date .
OpenScienceSymposium	conf:takesPlaceAt	dbpedia:Chemnitz .

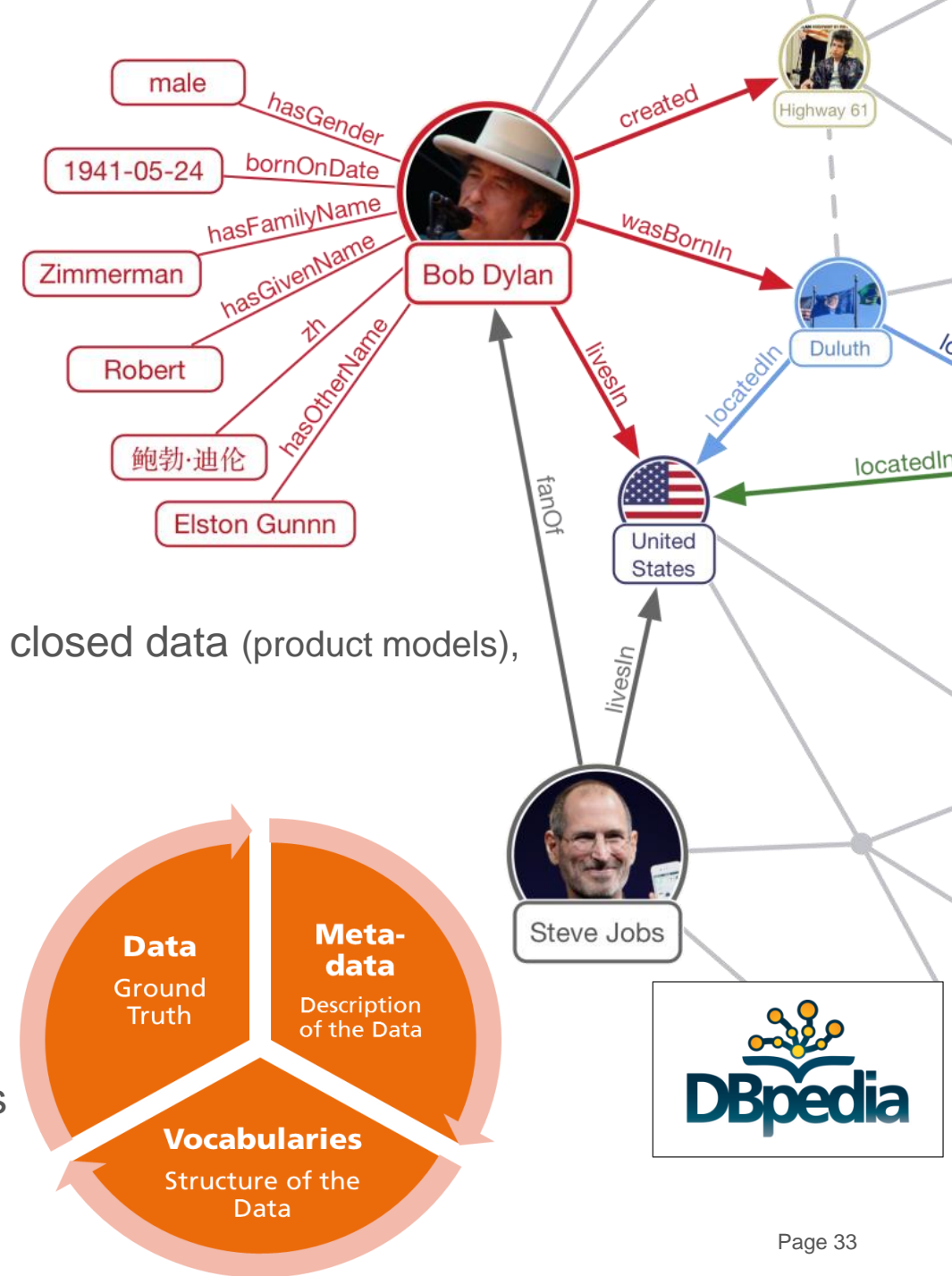
3. Publication under URL in Web, Intranet, Extranet

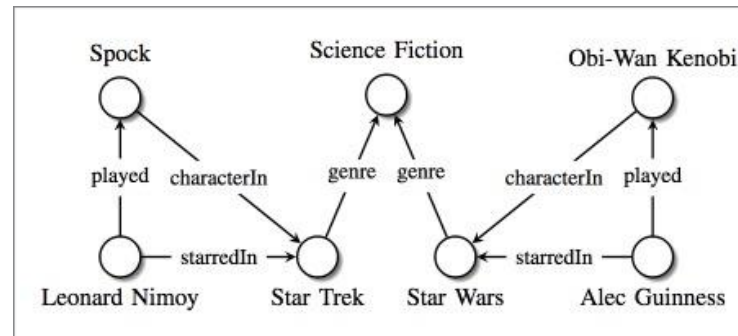
Creating Knowledge Graphs with RDF



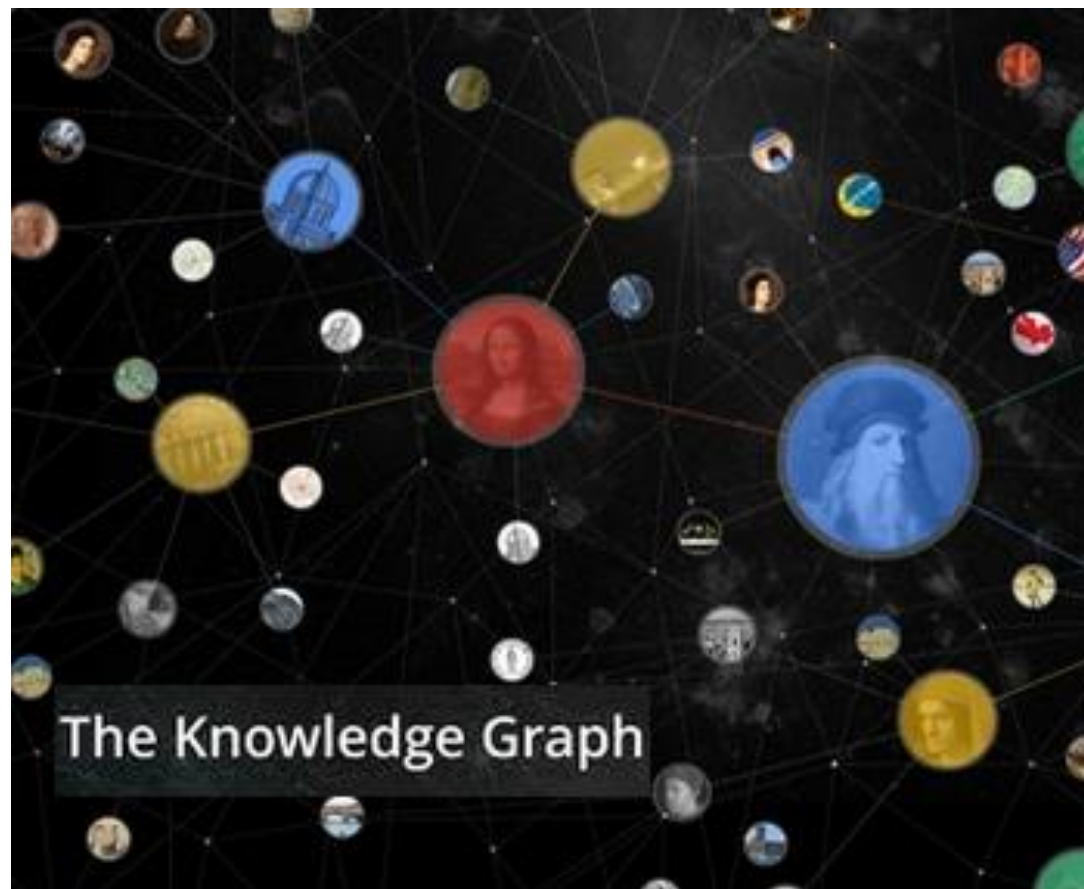
Knowledge Graphs – A definition

- Fabric of concept, class, property, relationships, entity desc.
- Uses a knowledge representation formalism (RDF, OWL)
- Holistic knowledge (multi-domain, source, granularity):
 - **instance data** (ground truth),
 - open (e.g. DBpedia, WikiData), private (e.g. supply chain data), closed data (product models),
 - derived, aggregated data,
 - **schema data** (vocabularies, ontologies)
 - **meta-data** (e.g. provenance, versioning, documentation licensing)
 - comprehensive **taxonomies** to categorize entities
 - **links** between internal and external data
 - **mappings** to data stored in other systems and databases





Source:
https://pic2.zhimg.com/v2-878ad2a55c440b18c889394a7abaa5d3_1200x500.jpg





GND	
Link zu diesem Datensatz	http://d-nb.info/gnd/1021356255
Typ	Person (piz)
Person	Appelbaum, Anne
Geschlecht	weiblich
Zeit	Lebensdaten: 1908-1998
Land	Deutschland (XA-DE); USA (XD-US)
Geografischer Bezug	Wirkungsort: New York, NY
Beruf(e)	Psychologin
Beziehungen zu Personen	Cassirer, Ernst (Vater)
Beteiligt an	1 Publikation 1. <i>Nachgelassene Manuskripte und Texte / Bd. 1. Zur Metaphysik der symbolischen Formen</i> 1995




Chemistry Example: CRISPR Genome Editing





THE PREPRINT SERVER FOR BIOLOGY

HOME | ABOUT | SUBMIT | ALERTS / RSS
| CHANNELS




Advanced Search

New Results

Previous

Next

A practical guide to CRISPR/Cas9 genome editing in Lepidoptera

Linlin Zhang,  Robert Reed


doi: <https://doi.org/10.1101/130344>

Now published in *Diversity and Evolution of Butterfly Wing Patterns* doi: [10.1007/978-981-10-4956-9_8](https://doi.org/10.1007/978-981-10-4956-9_8)

Abstract


Info/History


Metrics


 Preview PDF


Abstract


CRISPR/Cas9 genome editing has revolutionized functional genetic work in many organisms and is having an especially strong impact in emerging model systems. Here we summarize recent advances in applying CRISPR/Cas9 methods in Lepidoptera, with a focus on providing practical advice on the entire process of genome editing from experimental design through to genotyping. We also describe successful targeted GFP


 Download PDF


 Share

 Email

 Citation Tools

 Tweet

 Like 0



Subject Area

Genetics

Subject Areas

All Articles

Animal Behavior and Cognition

Chemistry Example: Populating the Graph

2. Adaptive Graph Curation & Completion



Author Robert Reed

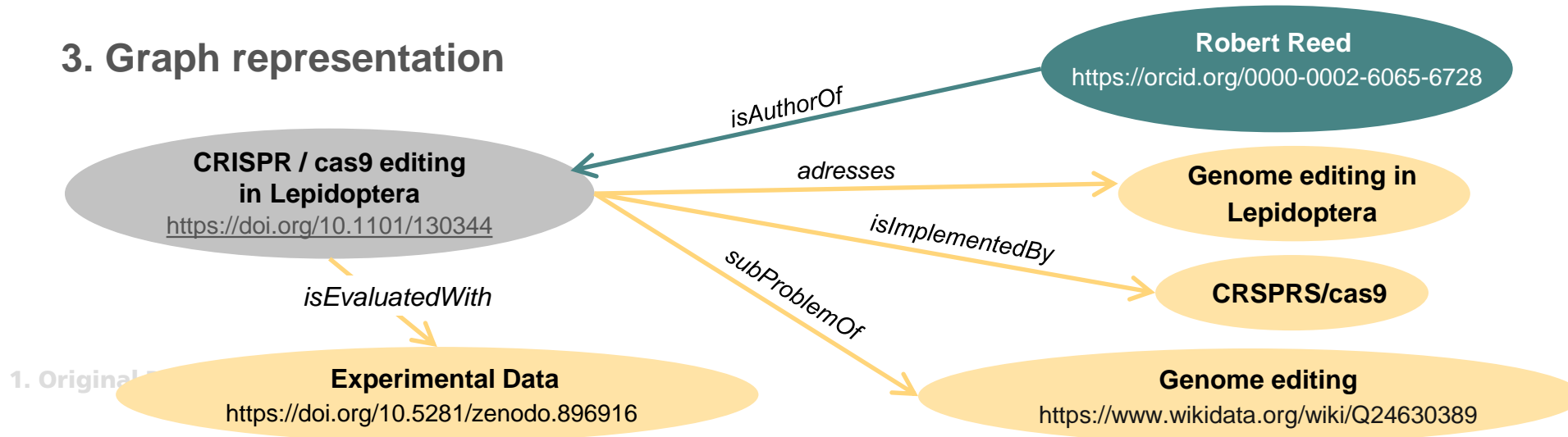
Research Problem Genome editing in Lepidoptera

Methods CRISPR / cas9

Applied on Lepidoptera

Experimental Data <https://doi.org/10.5281/zenodo.896916>

3. Graph representation



1. Original

Exploration and Question Answering

Research Challenge:

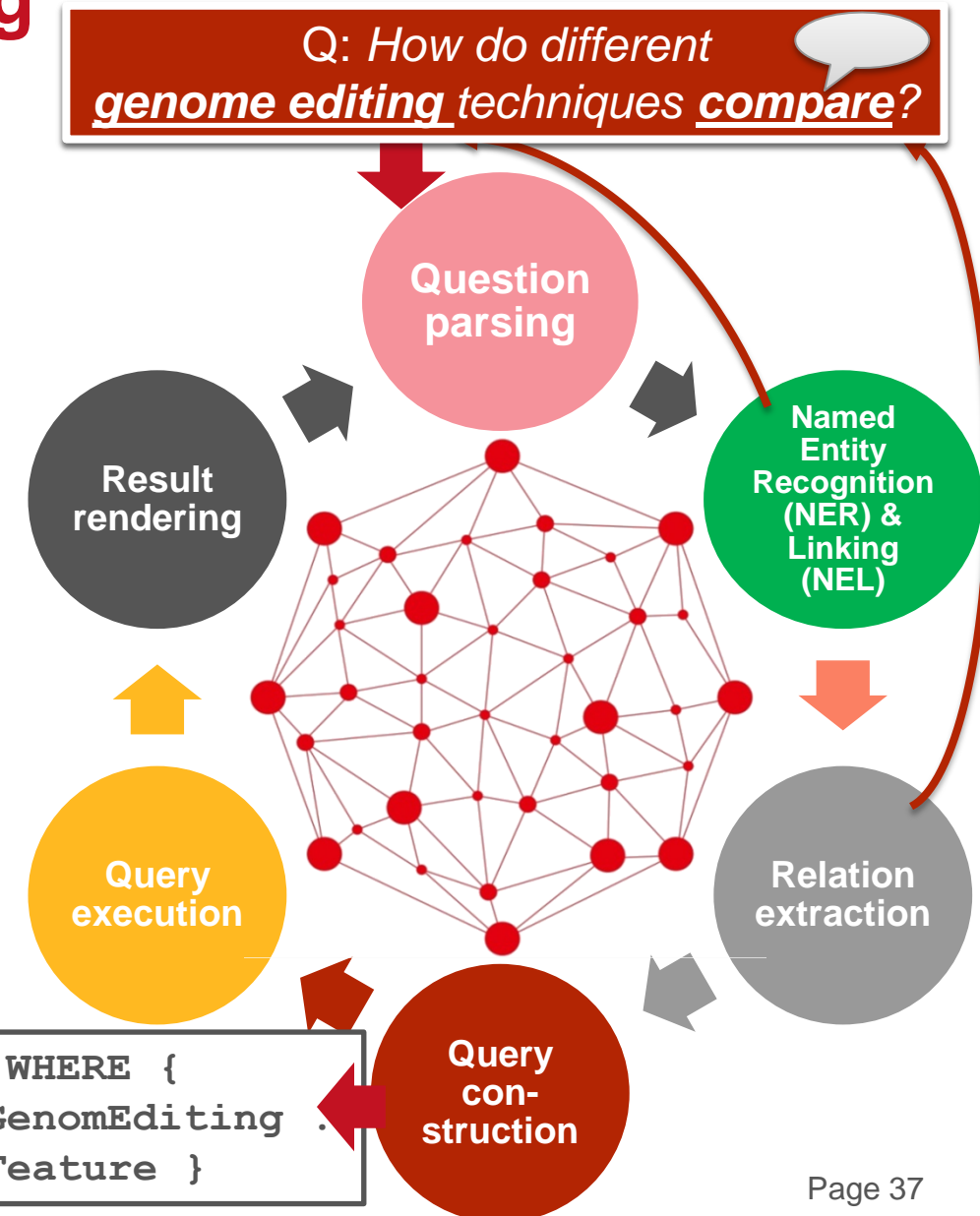
- Intuitive exploration leveraging the rich semantic representations
- Answer natural language questions

ScienceGRAPH Approach:

- KG-based QA component integration for dynamic and automated composition of QA pipelines for cognitive knowledge graphs (e.g. following [1])
- Round-trip refinement and integration of search, faceted exploration, question answering and conversational interfaces

[1] K. Singh, S. Auer et al: *Why Reinvent the Wheel? Let's Build Question Answering Systems Together.* The Web Conference (WWW 2018).

```
SELECT Approach, Feature WHERE {  
  Approach addresses GenomEditing .  
  Approach hasFeature Feature }
```



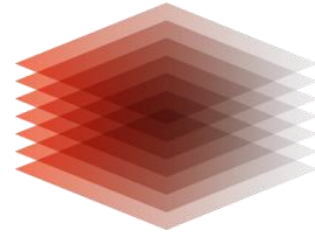
Result:

Automatic Generation of Comparisons / Surveys

Q: How do different **genome editing** techniques **compare**?

Engineered Nucleases	Site-specificity	Safety	Ease-of-use / costs/ speed
zinc finger nucleases (ZFN)	++ 9-18nt	+	-- \$\$\$: screening, testing to define efficiency
transcription activator-like effector nucleases (TALENs)	+++ 9-16nt	++	++ Easy to engineer 1 week / few hundred dollar
engineered meganucleases	+++ 12-40 nt	0	-- \$\$\$ Protein engineering, high-throughput screening
CRISPR system/cas9	++ 5-12 nt	-	+++ Easy to engineer few days / less 200 dollar

LEIBNIZ INFORMATION CENTRE
FOR SCIENCE AND TECHNOLOGY
UNIVERSITY LIBRARY



TIB

Demo: Open Research Knowledge Graph Prototype



[View all papers](#)

[Debug](#) ▼

[Add paper](#)



Homepage

Warning: the ORKG is currently in an alpha stage. Data you enter in the system can be deleted without any notice.

i The **Open Research Knowledge Graph** - or - ORKG aims to describe research papers and contributions in a structured manner. With ORKG research contributions become findable and comparable. In order to add your own research, or to contribute, [learn more](#) ↗

★ Browse by research field

Main

Arts and Humanities

Social and Behavioral
Sciences

Engineering

Physical Sciences &
Mathematics

Life Sciences

+ Recently added papers

A semi-automated, KNIME-based workflow for biofilm assays

Katrin Leinweber

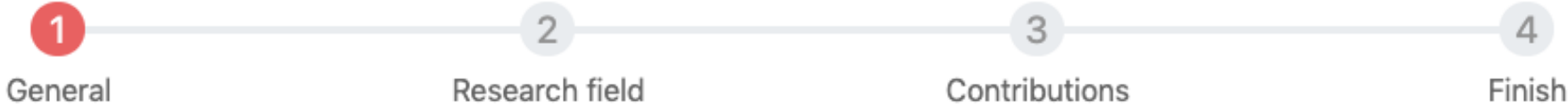
Open Research Knowledge Graph: Towards Machine Actionability in Scholarly Communication

Mohamad Yaser Jaradeh

Capsules of the diatom *Achnanthes minutissima* arise from fibrillar precursors and foster attachment of bacteria

Katrin Leinweber

Add paper



General paper data

By DOI

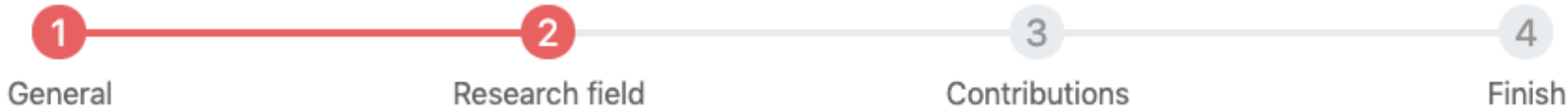
Manually

Paper DOI 

10.1147/rd.225.0509

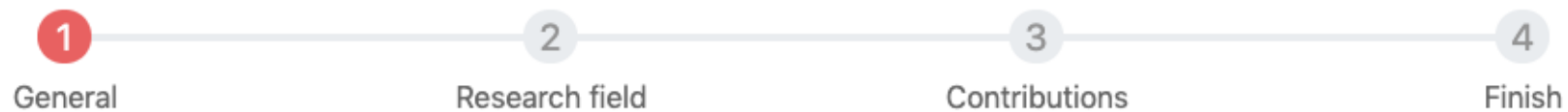
Lookup

Next step



Select the research field

Arts and Humanities	Statistics and Probability	Theory/Algorithms
Social and Behavioral Sciences	Physics	Software Engineering
Engineering	Oceanography and Atmospheric Sciences and Meteorology	OS/Networks
Physical Sciences & Mathematics	Mathematics	Programming Languages/Compilers
Life Sciences	Earth Sciences	Numerical Analysis/Scientific Computing
	Computer Sciences	



General paper data

By DOI

Manually

Paper DOI ?

10.1147/rd.225.0509

Lookup

Lookup result

Paper title: Algorithm and Hardware for a Merge Sort Using Multiple Processors

Authors: S. Todd

Publication date: April 2010

Next step



Specify research contributions

Contribution 1



+ Add another contribution

Research problems ?

Sorting algorithms ✕

Contribution data ?

← Back

Main

Merge sort

Programming language: C++



Stable: Yes



Best complexity: $n \log n$



Worst complexity: $n \log n$



+ Add property

Previous step

Next step

Algorithm and Hardware for a Merge Sort Using Multiple Processors

📅 April 2010

☰ Theory/Algorithms

👤 S. Todd

Contribution 1

Research problems

Sorting algorithms

Contribution data

← Back

Mair Merge sort

Worst complexity: $n \log n$



Best complexity: $n \log n$



Programming language: C++



Stable: Yes



Similar contributions [Show full comparison](#)

80% A variant of heapsort with almost optimal number of comparisons

54% Bubble sort: an archaeological algorithmic analysis

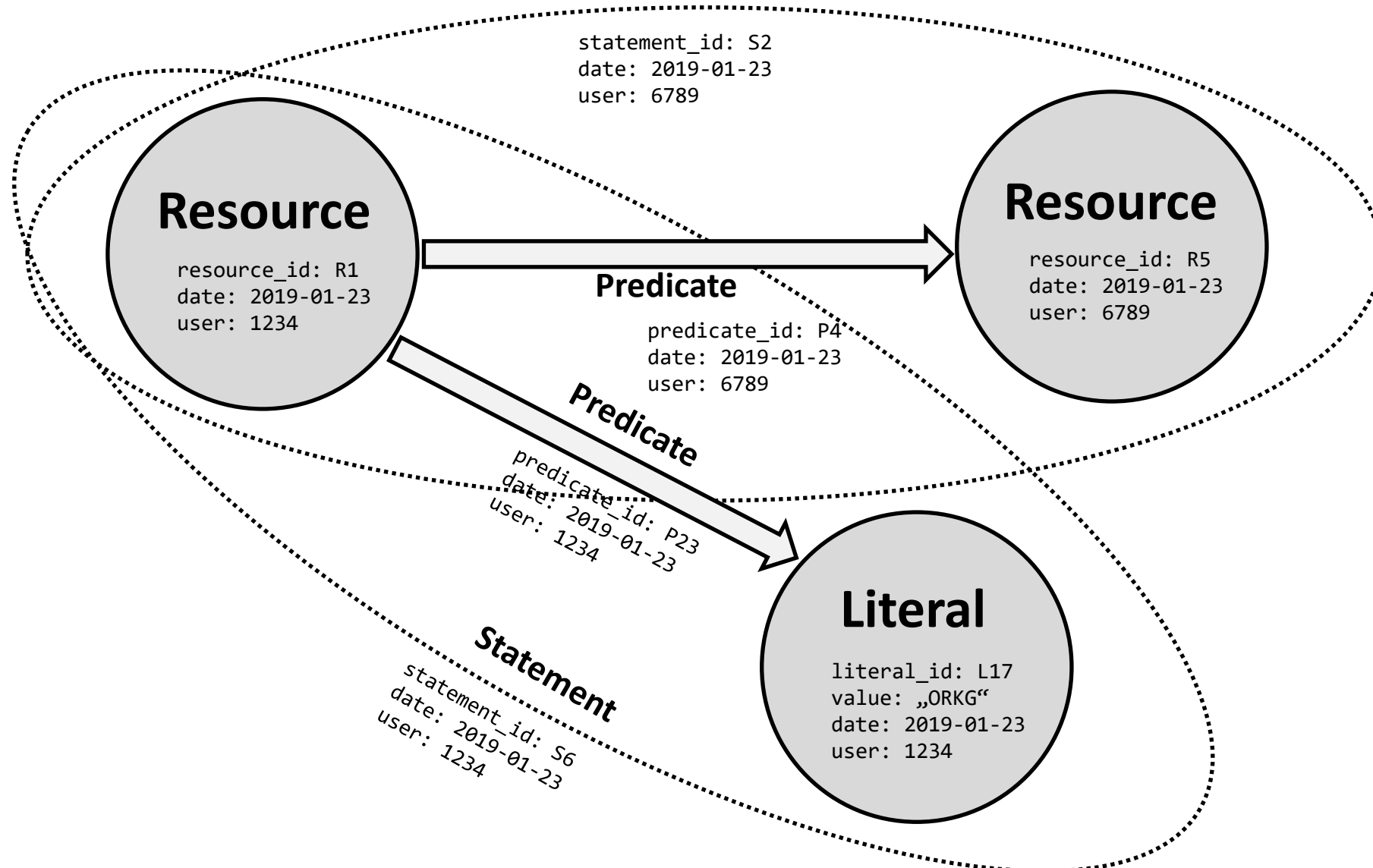
Compare

Add to comparison

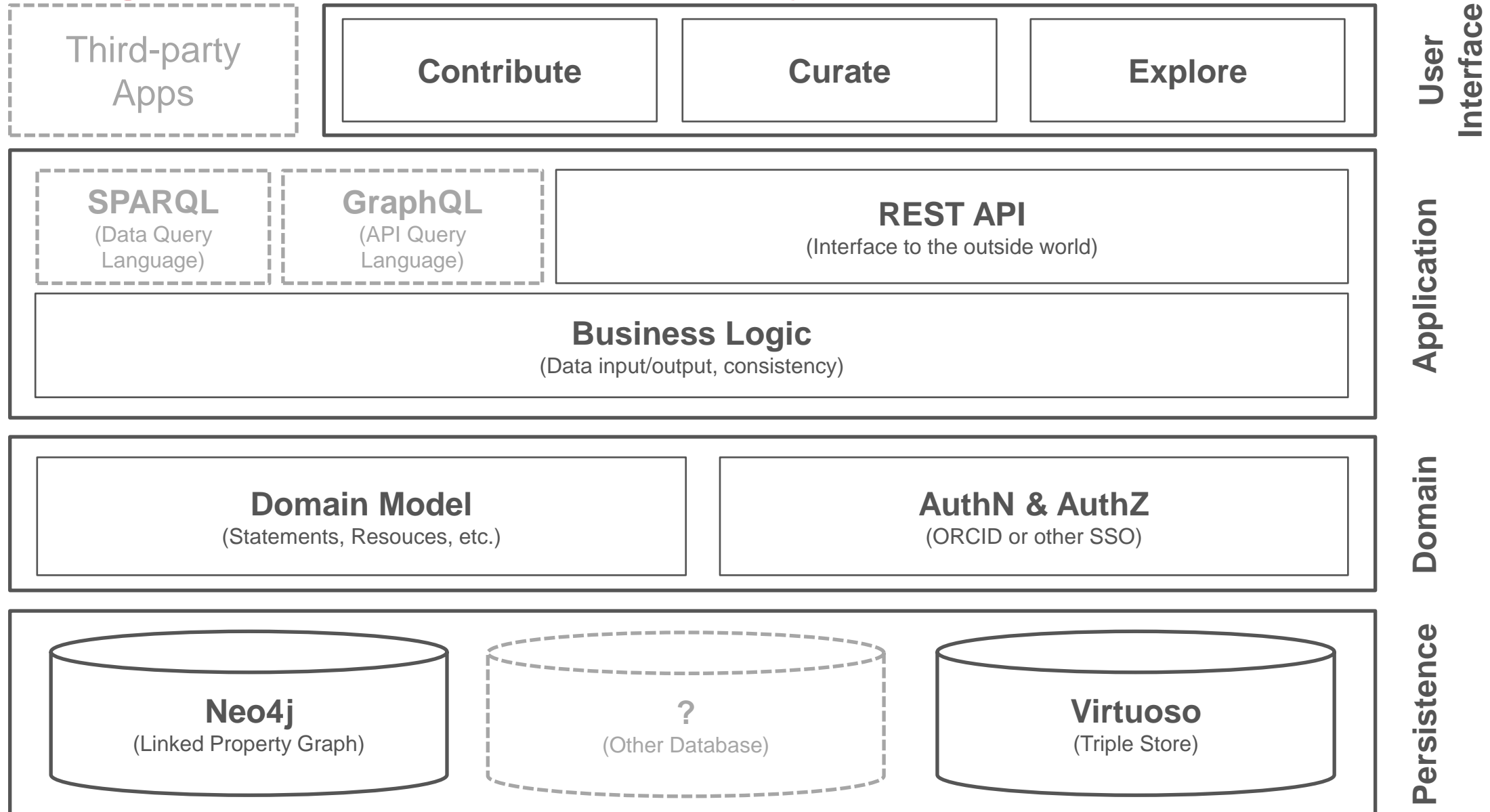
Properties	Algorithm and hardware for a merge sort using multiple processors Contribution #2	A variant of heapsort with almost optimal number of comparisons Contribution #1	Bubble sort: an archaeological algorithmic analysis Contribution #1
Algorithm	Merge sort	Heap sort	Bubble sort
Problem	Efficient sorting	Efficient sorting	Sorting
Programming language	C++	Empty	Python
Stable	Y	N	N
Best complexity	$n \log n$	n	n
Worst complexity	$n \log n$	$n \log n$	$n \log n$

High-level Data Model: RDF + Metadata

Statement



High-Level Architecture: Neo4j Graph Application



Iron-regulatory proteins secure iron availability in cardiomyocytes to prevent heart failure

**Saba Haddad^{1,2}, Yong Wang^{1,2}, Bruno Galy^{3,4}, Mortimer Korf-Klingebiel^{1,2},
Valentin Hirsch^{1,2}, Abdul M. Baru^{1,2}, Fatemeh Rostami^{1,2}, Marc R. Reboll^{1,2},
Jörg Heineke², Ulrich Flögel⁵, Stephanie Groos⁶, André Renner⁷, Karl Toischer⁸,
Fabian Zimmermann⁹, Stefan Engeli¹⁰, Jens Jordan¹⁰, Johann Bauersachs²,
Matthias W. Hentze³, Kai C. Wollert^{1,2}, and Tibor Kempf^{1,2*}**

¹Division of Molecular and Translational Cardiology, Hannover Medical School, Carl-Neuberg-Straße 1, 30625 Hannover, Germany; ²Department of Cardiology and Angiology, Hannover Medical School, Carl-Neuberg-Straße 1, 30625 Hannover, Germany; ³European Molecular Biology Laboratory, Meyerhofstraße 1, 69117 Heidelberg, Germany; ⁴Division of Virus-associated Carcinogenesis, German Cancer Research Centre, Im Neuenheimer Feld 280, 69120 Heidelberg, Germany; ⁵Department of Molecular Cardiology, University of Düsseldorf, Universitätsstraße 1, 40225 Düsseldorf, Germany; ⁶Institute of Cell Biology, Hannover Medical School, Carl-Neuberg-Straße 1, 30625 Hannover, Germany; ⁷Department of Thoracic and Cardiovascular Surgery, University of Bochum, Georgstraße 11, 32545 Bad Oeynhausen, Germany; ⁸Department of Cardiology and Pneumology, University of Göttingen, Robert-Koch-Straße 40, 37075 Göttingen, Germany; ⁹Department of Analytical Chemistry, Leibniz University Hannover, Callinstraße 1, 30167 Hannover, Germany; and ¹⁰Institute of Clinical Pharmacology, Hannover Medical School, Carl-Neuberg-Straße 1, 30625 Hannover, Germany

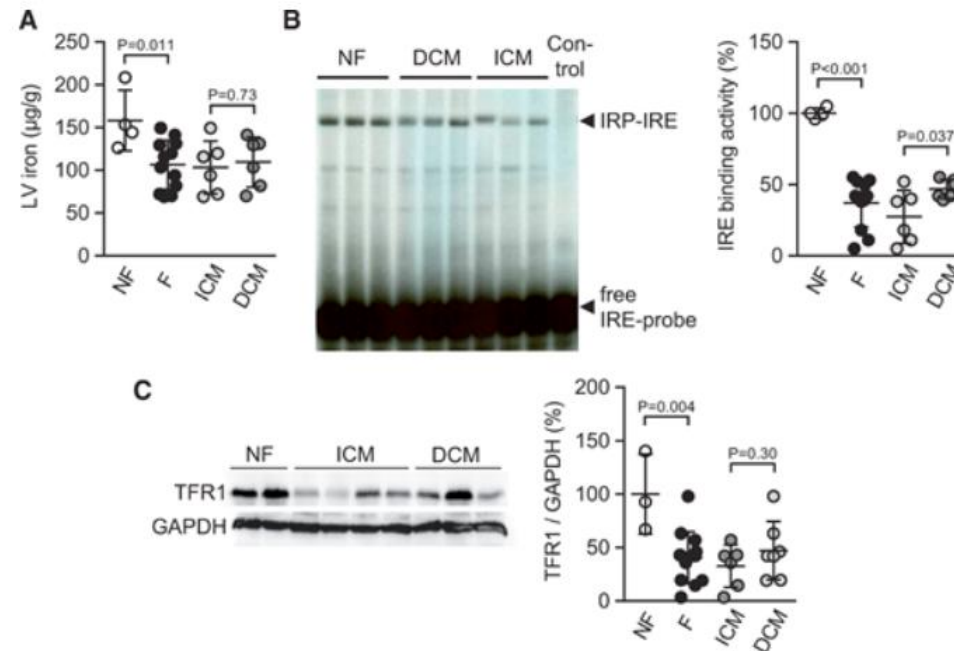
Received 30 November 2015; revised 27 June 2016; accepted 12 July 2016; online publish-ahead-of-print 21 August 2016

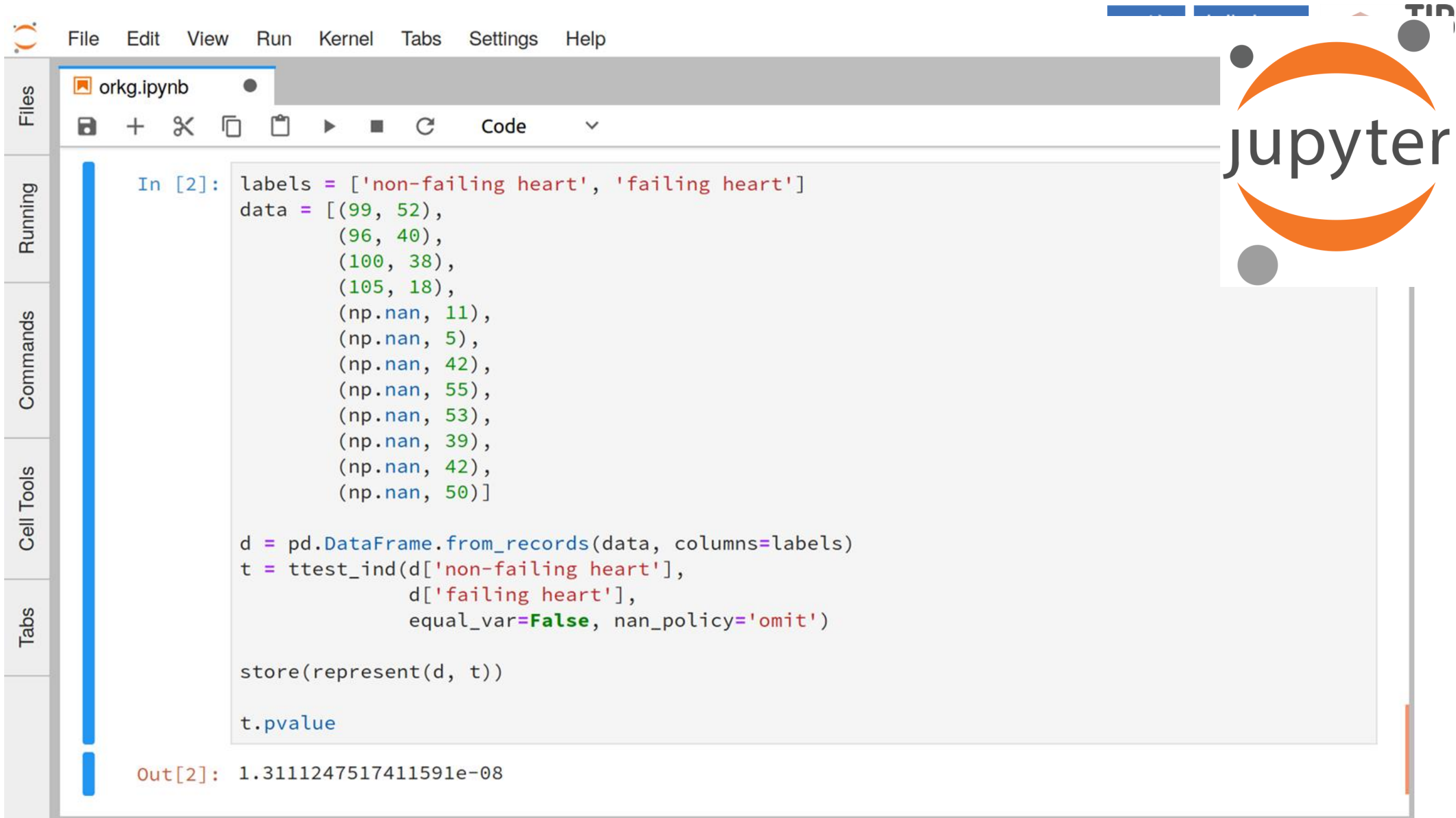
See page 373 for the editorial comment on this article (doi: 10.1093/eurheartj/ehw386)

Reduced iron content, IRE binding activity, and transferrin receptor expression in the failing human heart

Consistent with previous reports,^{5,6} iron concentration was significantly lower in LV tissue samples from patients with advanced heart failure than in LV tissue samples from unused donor hearts (Figure 1A). As shown by electrophoretic mobility shift assays, IRE binding activity was significantly reduced in failing hearts (most pronounced in patients with ischemic cardiomyopathy) (Figure 1B). Protein expression levels of the transferrin receptor were significantly lower in failing hearts than in the controls (Figure 1C).

Figure 1





The image shows a JupyterLab interface. On the left is a sidebar with tabs for 'Files', 'Running', 'Commands', 'Cell Tools', and 'Tabs'. The main area displays a code cell with the following Python code:

```
In [2]: labels = ['non-failing heart', 'failing heart']
data = [(99, 52),
        (96, 40),
        (100, 38),
        (105, 18),
        (np.nan, 11),
        (np.nan, 5),
        (np.nan, 42),
        (np.nan, 55),
        (np.nan, 53),
        (np.nan, 39),
        (np.nan, 42),
        (np.nan, 50)]

d = pd.DataFrame.from_records(data, columns=labels)
t = ttest_ind(d['non-failing heart'],
              d['failing heart'],
              equal_var=False, nan_policy='omit')

store(represent(d, t))

t.pvalue
```

Below the code cell, the output is shown:

```
Out[2]: 1.3111247517411591e-08
```

A large, semi-transparent Jupyter logo watermark is visible on the right side of the image.

Add paper

1

General

2

Research field

3

Contributions

4

Finish

General paper data

☒ By DOI☐ ManuallyPaper DOI [Lookup](#)[Next step](#)



General paper data

By DOI ☐ Manually ☐

Paper DOI 

10.1093/eurheartj/ehw333

Lookup

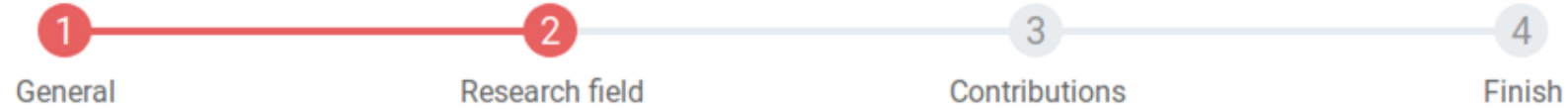
Lookup result

Paper title: Iron-regulatory proteins secure iron availability in cardiomyocytes to prevent heart failure

Authors: Saba Haddad, Yong Wang, Bruno Galy, Mortimer Korf-Klingebiel, Valentin Hirsch, Abdul M. Baru, Fatemeh Rostami, Marc R. Reboll, Jörg Heineke, Ulrich Flögel, Stephanie Groos, André Renner, Karl Toischer, Fabian Zimmermann, Stefan Engeli, Jens Jordan, Johann Bauersachs, Matthias W. Hentze, Kai C. Wollert, Tibor Kempf

Publication date: August 2016

Next step



Select the research field

Arts and Humanities	Nutrition	Systems and Integrative Physiology
Social and Behavioral Sciences	Forestry and Forest Sciences	Exercise Physiology
Engineering	Entomology Food Science	Endocrinology
Physical Sciences & Mathematics	Animal Sciences	Comparative and Evolutionary Physiology
Life Sciences	Physiology	Cellular and Molecular Physiology
	Nursing Pharmacology, Toxicology and Environmental Health	

[Previous step](#)[Next step](#)



Specify research contributions

Contribution 1

+ Add another contribution

Research problems ?

Iron deficiency in heart failure patients

Contribution data ?

No values

+ Add property

Contribution 1



+ Add another contribution

Research problems ?

Iron deficiency in heart failure patients ✕

Contribution data ?

Yields

 Delete

Object ▾

IRE

Cancel

Done

+ Add property

Statistically significant hypothesis test with IRE
binding dependent variable on failing and non-
failing hearts

Previous step

Next step

Contribution 1



+ Add another contribution

Research problems ?

Iron deficiency in heart failure patients ✕

Contribution data ?

Yields

 DeleteStatistically significant hypothesis test with IRE binding dependent variable on failing and non-failing hearts

+ Add value

+ Add property

Previous step

Next step

Research problems

Iron deficiency in heart failure patients

Contribution data

← Back

Ma Statistically significant hypothesis test with IRE binding dependent variable on failin

Has specified output: *the p-value of the statistical hypothesis test*



Type: *two sample t-test with unequal variance*



Has specified input: *3 values*



Label: *Statistically significant hypothesis test with IRE binding dependent variable on failing and non-failing hearts*



Research problems

Iron deficiency in heart failure patients

Contribution data

← Back

Main

Statistics

the p-value of the statistical hypothesis test

Has value specification: *the value specification of the p-value*



Type: *p-value*



Label: *the p-value of the statistical hypothesis test*



Research problems

Iron deficiency in heart failure patients

Contribution data

← Back

Main

Stat

the p

the value specification of the p-value

Type: *scalar value specification*



Has specified numeric value: $1.3111247517411591e-08$



Label: *the value specification of the p-value*



Research problems

Iron deficiency in heart failure patients

Contribution data

← Back

Ma Statistically significant hypothesis test with IRE binding dependent variable on failin

Has specified output: *the p-value of the statistical hypothesis test*



Type: *two sample t-test with unequal variance*



Has specified input



failing heart

the study design dependent variable

non-failing heart

Label: *Statistically significant hypothesis test with IRE binding dependent variable on failing and non-failing hearts*



Research problems

Iron deficiency in heart failure patients

Contribution data

← Back

Main

Stat

the study design dependent variable

Type

study design dependent variable

iron-responsive element binding

Label: *the study design dependent variable*

Term of the Gene Ontology, namely GO:0030350

Research problems

Iron deficiency in heart failure patients

Contribution data

← Back

Main

Stat

non-failing heart

Has part

the scalar measurement datum #3

the scalar measurement datum #1

the scalar measurement datum #4

the scalar measurement datum #2

Type: *continuous variable*

Label: *non-failing heart*

Research problems

Iron deficiency in heart failure patients

Contribution data

← Back

Main

Statistics

non-

the s

the value specification #4

Type: *scalar value specification*



Has specified numeric value: *105.0*



Label: *the value specification #4*



Research problems

Iron deficiency in heart failure patients

Contribution data

← Back

Main

Statistics

non-

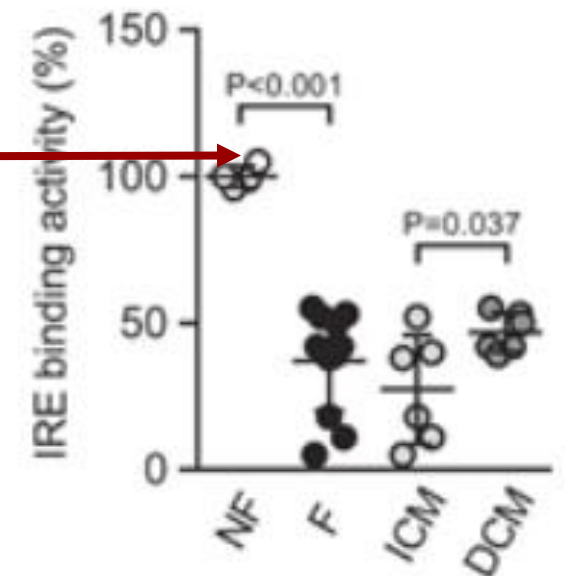
the s

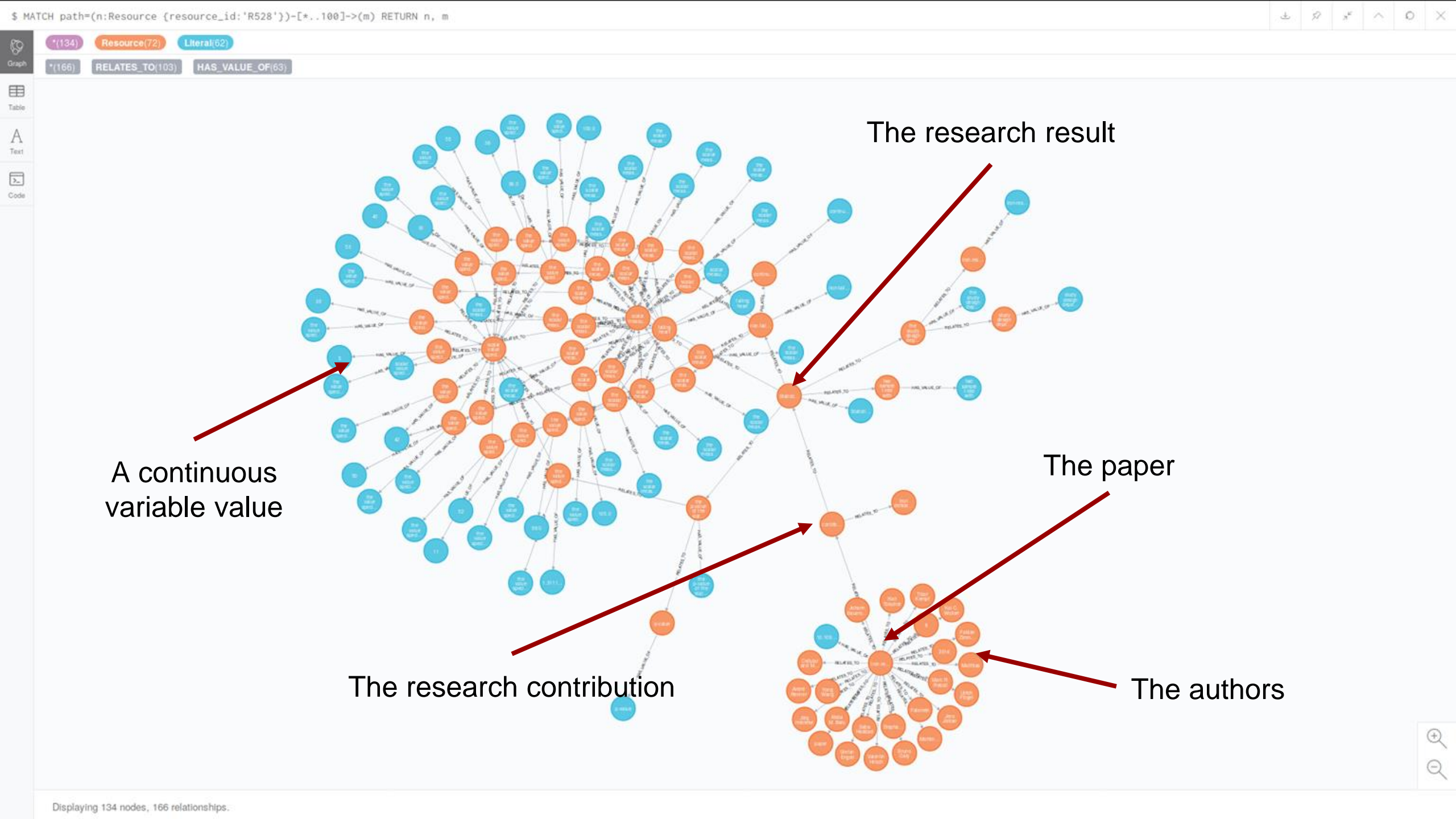
the value specification #4

Type: *scalar value specification*

Has specified numeric value: *105.0*

Label: *the value specification #4*





The Team



Group Leaders



Dr. Markus Stocker



Dr. Gábor Kismihók

PostDocs

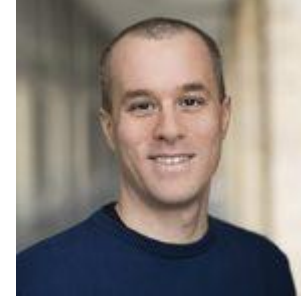


Dr. Javad Chamanara



Dr. Jennifer D'Souza

Software Development



Manuel Prinz



Wazed Ali

Doctoral Researchers



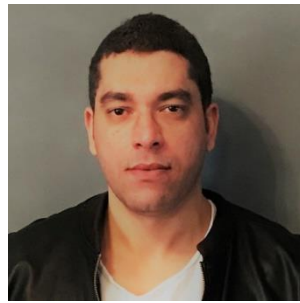
Olga Lezhnina



Allard Oelen



Yaser Jaradeh



Shereif Eid



Sarven Capadisli



Vitalis Wiens



Collaborators TIB/L3S Scientific Data Management



Prof. (Univ. S. Bolivar)
Dr. Maria Esther Vidal



Kemele Endris



Farah Karim

Collaborators InfAI Leipzig / AKSW



Dr. Michael Martin



Natanael Arndt

Project Management



Alex Garatzogianni



Laura Granzow

Conclusions

- We need to **reinvent scholarly communication**
- **Knowledge Graphs** are perfectly suited to **capture research contributions** in a **structured and semantic way** making them **human and machine interpretable**
- With our **Open Research Knowledge Graph initiative** we aim to **establish a registry for research contributions** (maybe similar to some extent as Crossref for DOIs/bibliographic metadata)

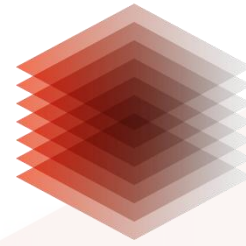
Stay tuned

- <https://tib.eu>
- Mailinglist/group:
<https://groups.google.com/forum/#!forum/orkg>
- Open Research Knowledge Graph:
<https://orkg.org>
- ERC Consolidator Grant ScienceGRAPH started in May



European
Research
Council

LEIBNIZ INFORMATION CENTRE
FOR SCIENCE AND TECHNOLOGY
UNIVERSITY LIBRARY



TIB



<https://de.linkedin.com/in/soerenauer>



<https://twitter.com/soerenauer>



https://www.xing.com/profile/Soeren_Auer



http://www.researchgate.net/profile/Soeren_Auer

Contact

Prof. Dr. Sören Auer
TIB & Leibniz University of Hannover
auer@tib.eu



Leibniz
Universität
Hannover

Leibniz
Leibniz
Association